



## Verità e paradossi: una rilettura dell' *Outline* di Kripke

a cura di Alessandro Becchi e Fulvio Guatelli

*La presente rubrica si propone di dare un'esposizione contestualizzata di un particolare problema filosofico, attraverso la rilettura di un articolo, saggio, o libro particolarmente influente e ricco di sviluppi per l'argomento in questione. Essa non vuole (né potrebbe) essere in alcun modo esaustiva degli argomenti trattati; ciò che qui intendiamo offrire è semplicemente uno strumento che consenta al lettore di approfondire una tematica filosofica eminente. A tale scopo includiamo una breve bibliografia ragionata al termine della presentazione.*

L'ambito di ricerca che intendiamo porre all'attenzione del lettore concerne la questione della *definibilità* della verità - che cosa significa che qualcosa è vero? - tralasciando il problema epistemologico del criterio della verità: come conosciamo che qualcosa è vero? La verità sarà considerata in questa sede nella sua accezione *semantica* ('vero' come predicato di enunciati, proposizioni, o proferimenti), e non nell'accezione *ontologica* che tale concetto ha assunto tanto in gran parte della storia della filosofia, quanto nel linguaggio naturale ('vero' come sinonimo di 'reale', 'autentico').

Alfred Tarski, nel saggio *The Concept of Truth in Formalized Languages*,<sup>1</sup> ha elaborato una teoria del predicato 'vero' non contraddittoria ('formalmente corretta') e capace di catturare una caratteristica semantica fondamentale del

<sup>1</sup> In A. Tarski, *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*, Oxford, Clarendon Press, 1956, pp. 152-277 (vers. orig., in polacco, 1933); trad. it. in F. Rivetti Barbò, *L'antinomia del mentitore nel pensiero contemporaneo. Da Peirce a Tarski*, Milano, Vita e pensiero, 1961, pp. 391-677.

concetto intuitivo di verità: il fatto, cioè, che l'asserire la verità di un enunciato equivale ad asserire l'enunciato stesso ('materialmente adeguata'). Duplice era l'obiettivo di Tarski, il primo, legittimare le nozioni semantiche all'interno di una 'concezione scientifica del mondo' stando alla quale la nozione di verità appariva sospetta di metafisica: a tale proposito egli mostrò come ridurre tale nozione a concetti logico-matematici; il secondo, eliminare i paradossi semantici. A questo proposito Tarski ha mostrato [Teorema di Tarski, 1933] che non può esistere un linguaggio interpretato  $L$  tale che:

(i) sia *semanticamente chiuso*, ossia contenga un nome per ogni suo enunciato, un predicato 'vero-in- $L$ ' per enunciati, e disponga di risorse sintattiche sufficienti a innescare il meccanismo dell'*autoriferimento*: cioè sia in grado di formulare almeno un'istanza del *mentitore*;<sup>2</sup>

(i) sia tale che il predicato 'vero-in- $L$ ' soddisfi la condizione di *adeguatezza materiale*: per ogni enunciato  $A$  di  $L$ , e per ogni ' $A$ ' nome di  $A$ , valga: ' $A$  è vero-in- $L$  se e solo se  $A$ ;

(ii) valga la *logica bivalente* classica: ogni enunciato di  $L$  ha uno e uno solo dei seguenti valori di verità, il vero e il falso.

Di fronte a questo fatto, la soluzione adottata da Tarski per definire 'vero-in- $L$ ' all'interno di una teoria consistente è quella di rifiutare (i), ossia escludere la possibilità che  $L$  sia semanticamente chiuso. La conseguenza di ciò consiste nel definire 'vero-in- $L$ ' in un metalinguaggio  $L_1$  più potente di  $L$ , ed eventualmente, definire 'vero-in- $L_1$ ' in un meta-metalinguaggio  $L_2$  più potente di  $L_1$ , e così via; ne risulta una gerarchia (potenzialmente) infinita di linguaggi nella quale, ad ogni livello, viene definita la verità per gli enunciati del livello precedente; ciò comporta la definizione di un *diverso* predicato di verità per ogni linguaggio oggetto considerato. In questo contesto i paradossi semantici (mentitore) non risultano più esprimibili in quanto sintatticamente scorretti.

Nel 1975 Saul Kripke, con la pubblicazione del saggio *Outline of a Theory of Truth*, ha dato il via a un'ampia serie di ricerche su possibili soluzioni non tarskiane ai paradossi semantici e su teorie della verità alternative. Alla base del

<sup>2</sup> Una possibile formulazione del 'mentitore': sia dato un linguaggio  $L$  contenente nomi per ogni suo enunciato, un predicato di verità  $T(x)$ , e un nome  $t$  che denoti  $\neg T(t)$ . Supponiamo che  $T(t)$  valga; abbiamo, per la logica dell'identità,  $T(\neg T(t))$ , dal momento che  $t = \neg T(t)$ , da cui per  $T$ -equivalenza:  $\neg T(t)$ , contraddizione. Supponiamo allora che valga  $\neg T(t)$ ; per  $T$ -equivalenza, abbiamo  $T(\neg T(t))$ , e, sempre per la logica dell'identità,  $T(t)$ , contraddizione.

lavoro di Kripke c'è l'idea che è possibile costruire una teoria della verità che (i) contiene un *unico* predicato 'vero' (applicabile ad enunciati contenenti il predicato stesso) e (ii) nella quale i paradossi e gli enunciati infondati vengono spiegati attraverso lacune di valore di verità [*truth-value gaps*].

*La diagnosi dei paradossi:* partendo da un esempio paradigmatico Kripke mostra che non c'è un criterio intrinseco (generale) che ci metta in grado di scartare – come insignificanti o mal formati – quegli enunciati che portano a paradosso; l'enunciato “La maggior parte delle asserzioni di Nixon sul Watergate sono false”, chiaramente ben formato e significante, dà luogo a paradosso sotto determinate assunzioni (Kripke ne fornisce un argomento nelle pp. 691-692). Risulta dunque che la maggior parte delle nostre asserzioni ordinarie riguardo alla verità / falsità sono soggette – se le circostanze empiriche sono estremamente sfavorevoli – ad esibire caratteristiche paradossali. Questa tesi implica che l'analisi tarskiana dei paradossi non è del tutto adeguata: nel quadro della teoria di Tarski gli enunciati paradossali sono infatti mal formati. Supponiamo che Dean asserisca (1) “Tutti i proferimenti di Nixon sul Watergate sono falsi”, mentre Nixon asserisce: (2) “Tutti i proferimenti di Dean sul Watergate sono falsi”. Chiaramente ognuno dei due enunciati parla anche dell'altro; stando all'approccio tarskiano, però, risulta che ciò è impossibile: infatti (i) se (1) e (2) sono sullo stesso livello allora non possono parlare l'uno dell'altro (altrimenti uno dei due sarebbe mal formato), e (ii) se sono su livelli diversi, il più alto parlerà del più basso ma non viceversa, tradendo il significato intuitivo dei due enunciati. Tuttavia, osserva Kripke, possiamo assegnare valori di verità [v.d.v.] non ambigui a (1) e (2), malgrado parlino l'uno dell'altro, sotto opportune ipotesi empiriche (pp. 695-697).

Vi sono anche enunciati del tipo del *veritiero*: ‘Questo enunciato è vero’ che, pur non portando a paradosso, non sono suscettibili di valutazione e quindi sono infondati; Kripke caratterizza un enunciato come *fondato* se il processo mediante cui lo valutiamo ci porta in ultima analisi ad enunciati non contenenti il concetto di verità (falsità), altrimenti *infondato*. Egli fa inoltre notare come, parallelamente al concetto di paradossalità, la fondatezza e l'infondatezza non sono – in generale – caratteristiche intrinseche di un enunciato, ma possono dipendere dalle circostanze empiriche.

*Descrizione informale della teoria:* delle caratteristiche di L (indicate nella diagnosi tarskiana) da cui scaturiscono i paradossi, Kripke rifiuta il punto (iii), ossia la bivalenza; adotta quindi una logica trivalente ('Kleene forte') che ammette, accanto al vero e al falso, un terzo valore di verità, l'indeterminato. Dal momento che in tale contesto la negazione di un enunciato indeterminato è indeterminata, l'istanza del mentitore riportata in nota 1 non risulta più paradossale, poiché se  $T(t)$  è indeterminato, lo è anche  $\neg T(t)$ .

La teoria è costruita mediante una *gerarchia transfinita di interpretazioni* (modelli) del predicato 'vero',  $T(x)$ . Al livello più basso della gerarchia ( $M_0$ ) tale predicato è completamente indefinito: preso un qualsiasi enunciato (vero o falso che sia nell'interpretazione standard del linguaggio  $L$ ) non è possibile asserire a questo livello la sua verità o falsità. Passando al livello successivo ( $M_1$ ) il predicato 'vero' assume come propria *estensione* l'insieme degli enunciati veri in  $M_0$  e come *anti-estensione* l'insieme di quelli falsi in  $M_0$ , con la conseguenza che acquistano v.d.v. determinato nuovi enunciati (operazione di 'salto kripkeano'). Salendo nella gerarchia, il predicato di verità cresce sia in estensione sia in anti-estensione: in questo modo sono sempre meno gli enunciati che restano con v.d.v. indeterminato. Reiterando questo processo un numero sufficientemente alto di volte (passaggio al transfinito), arriviamo al *punto fisso minimale*  $M_s$ , in cui tutti gli enunciati suscettibili di valutazione acquistano v.d.v. determinato; i soli enunciati che restano indeterminati sono quelli infondati e quelli paradossali. Infatti nel punto fisso minimale vale che: per ogni  $A$  di  $L$ ,  $M_s$  soddisfa  $T(A)$  se e solo se  $M_s$  soddisfa  $A$ , e  $M_s$  non soddisfa  $T(A)$  se e solo se  $M_s$  non soddisfa  $A$  (condizione di adeguatezza materiale); ciò significa che tutti gli enunciati valutati veri / falsi nel punto fisso minimale soddisfano la nozione intuitiva di fondatezza.

Fin qui la teoria non distingue ancora tra enunciati paradossali (es. il mentitore) ed enunciati infondati non paradossali (es. il veritiero); si possono costruire punti fissi non minimali che rendono via via veri / falsi un numero sempre più ampio di enunciati infondati non paradossali senza che ciò renda la teoria inconsistente. Consideriamo ad esempio il veritiero 'Questo enunciato è vero' e stabiliamo *arbitrariamente* il suo valore di verità, poniamo vero; supponiamo allora di iniziare la gerarchia di interpretazioni di  $T(x)$  ponendo nella sua estensione tale enunciato, per poi continuare nel modo canonico al passo; in modo del tutto analogo al processo sopra descritto si raggiungerà il punto fisso non minimale  $M_d$  tale che  $M_d$  soddisfa 'Questo enunciato è vero'. Tale punto fisso attribuisce così v.d.v. ad un enunciato che nel punto fisso minimale era indeterminato. Si può dimostrare che ogni punto fisso può essere esteso ad un *punto fisso massimale*, vale a dire, un punto fisso che non ha estensioni proprie che siano punti fissi. Un punto fisso massimale assegna "tanti valori di verità quanti se ne possono assegnare", cioè assegna un valore di verità ad ogni enunciato infondato non paradossale. Si giunge così ad un tripartizione degli enunciati di  $L$  relativamente ad una interpretazione di partenza di tutti gli enunciati in cui non occorre 'vero':

- (a) un enunciato è *fondato* se e solo se è valutato vero / falso nel punto fisso minimale;
- (b) un enunciato è *paradossale* se e solo se non è valutato vero / falso in

alcun punto fisso;

(c) un enunciato è *infondato non paradossale* se e solo se è valutato vero / falso in almeno un punto fisso non minimale.

La teoria appena formulata dà conto del processo di *apprendimento* della parola 'vero' nel linguaggio naturale. L'apprendimento dell'uso del predicato 'vero' coinvolge sostanzialmente due capacità: (i) quella di asserire che un enunciato è vero esattamente nei casi in cui siamo disposti ad asserire l'enunciato stesso (T-equivalenze); (ii) quella di ridurre – attraverso la logica – enunciati in cui occorre il predicato 'vero' a T-equivalenze (es. p. 701). In modo analogo, il livello  $M_0$  corrisponde alla fase in cui il soggetto non ha alcuna comprensione di enunciati del tipo  $T(A)$ : infatti a questo livello  $T(A)$  è indeterminato; al livello  $M_1$  l'interpretazione di  $T(x)$  è costituita dai nomi degli enunciati veri/falsi di  $M_0$ , cioè l'enunciato  $T(A)$  vale sse vale  $A$  al livello  $M_0$  (T-equivalenze). Alla fine, quando il processo diviene saturato il soggetto raggiunge il punto fisso  $M_s$ , ossia è capace di valutare qualsiasi enunciato suscettibile di valutazione; in altri termini, egli ha afferrato la regola generale dell'uso del predicato 'vero'. Il più piccolo punto fisso  $M_s$  è da un lato, il modello più naturale per il nostro concetto intuitivo di verità, dall'altro, quello generato dalle nostre istruzioni per l'apprendimento del predicato 'vero'.

*Meriti dell'Outline.* Kripke ha dimostrato in modo rigoroso che esistono linguaggi capaci di autoriferimento in grado di definire il proprio predicato di verità. Dando per buona la legittimità della caratteristica dell'autoriferimento, ciò fornisce una possibile soluzione del dissidio manifestato dal Teorema di Tarski intorno a due fondamentali intuizioni semantiche: il *principio di bivalenza* e il *criterio di adeguatezza materiale*, dal momento che per gli enunciati fondati – gli enunciati la cui valutazione soddisfa le nostre attese in modo 'naturale' – valgono entrambi.

Questo risultato, che deve i propri strumenti di attuazione in modo significativo alle ricerche del decennio precedente, insieme all'originalità della tesi kripkeana della relatività rispetto ad un modello delle nozioni di enunciato fondato ed enunciato paradossale, e quindi dell'assenza di un criterio generale – semantico o sintattico – per discernarli, costituisce probabilmente il maggiore contributo a quell'ambito della ricerca filosofica che vede nello studio dei paradossi una via eminente, o addirittura privilegiata, alla delucidazione dei principi che regolano la *semantica ingenua*.

*Bibliografia:*

- S. A. Kripke, *Outline of a Theory of Truth*, in «Journal of Philosophy», 72 (1975), pp. 690-716; ristampato in Martin, 1984, pp. 53-81; trad. it. in G. Franci (a cura di), *Esistenza e necessità*, Firenze, Ponte alle Grazie, 1992, pp. 94-118.
- R. L. Martin (a cura di), *Recent Essays on Truth and the Liar Paradox*, London, New York, Oxford University Press, 1984.
- P. Minari, *La verità è definibile?*, in «Rivista di filosofia», 83 (1992), pp. 77-103.
- M. Sheard, *A Guide to Truth Predicates in the Modern Era*, in «Journal of Symbolic Logic», 59 (1994), pp. 1032- 1054.
- A. Visser, *Semantics and the Liar Paradox*, in D. Gabbay, F. Guenther (a cura di), *Handbook of Philosophical Logic*, Dordrecht, Reidel, 1989, vol. IV, pp. 617-706.

*Più ampi, completi e difficili:*

- A. Gupta, N. Belnap, *The Revision Theory of Truth*, Cambridge (Mass.), MIT Press, 1993.
- V. McGee, *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*, Indianapolis, Hackett, 1991.